



La prise en compte de l'environnement par les agriculteurs : une analyse avec le package "ClustOfVar"

Vanessa Kuentz-Simonet, Sandrine Lyser, Marie Chavent, Jérôme Saracco, Jacqueline Candau, Philippe Deuffic

► To cite this version:

Vanessa Kuentz-Simonet, Sandrine Lyser, Marie Chavent, Jérôme Saracco, Jacqueline Candau, et al..
La prise en compte de l'environnement par les agriculteurs : une analyse avec le package "ClustOfVar".
1ères Rencontres R, Jul 2012, Bordeaux, France. hal-00717496

HAL Id: hal-00717496

<https://hal.science/hal-00717496>

Submitted on 13 Jul 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

La prise en compte de l'environnement par les agriculteurs : une analyse avec le package “ClustOfVar”

V. Kuentz-Simonet^a, S. Lyser^a, M. Chavent^b, J. Saracco^b, J. Candau^a, and P. Deuffic^a

^a Irstea, UR ADBX,
50 avenue de Verdun, 33612 Cestas Cedex, France
{vanessa.kuentz-simonet,sandrine.lyser,jacqueline.candau,philippe.deuffic}@irstea.fr

^b INRIA Bordeaux Sud Ouest,
Institut de Mathématiques de Bordeaux,
Université de Bordeaux,
351 cours de la libération, 33405 Talence Cedex, France
{marie.chavent, jerome.saracco}@math.u-bordeaux1.fr

Mots clefs : classification de variables, variables synthétiques, perception environnementale, typologie.

En statistique exploratoire multidimensionnelle, la classification des observations est couramment utilisée pour établir des profils-types. Une stratégie classique consiste à réaliser une analyse factorielle des données puis à appliquer une méthode de classification sur les scores des individus mesurés sur les composantes principales obtenues. Cependant certains auteurs (De Soete et Carroll, 1994 ou Vichi et Kiers, 2001) ont souligné les effets néfastes de cette procédure dite « tandem analysis ». Dans le cas particulier de données quantitatives, ils montrent que l'ACP identifie parfois des composantes qui contribuent peu à la détection d'une structure dans les observations ou qui au contraire masquent l'information taxinomique. En effet, on peut concevoir que des informations relatives à la structure des observations puissent être masquées par la création de ces composantes orthogonales qui visent à reconstruire au mieux la variance contenue dans le nuage de points initial. L'approche par classification de variables proposée ici (Chavent et al. 2011) est une alternative à la première étape d'analyse factorielle pour la typologie d'observations. Cette méthode permet de réduire la dimension du tableau en supprimant l'information redondante. En réorganisant les variables en classes homogènes, elle conduit à construire simultanément des variables synthétiques sans imposer de contraintes d'orthogonalité. De plus, dans chaque classe de variables, les coordonnées des modalités des variables qualitatives la composant nous permettent de visualiser ces variables synthétiques comme une sorte de gradient. Il est alors possible d'interpréter et de labelliser ces variables synthétiques aisément. Ainsi la compréhension de la typologie des observations est simplifiée.

La méthodologie proposée est illustrée à partir d'une enquête réalisée en 2005 par des sociologues d'Irstea auprès d'agriculteurs français. L'objectif de cette étude était de “cerner la façon dont les agriculteurs conçoivent la protection de l'environnement en relation avec leur activité” (Candau et al., 2005). L'approche proposée permet de mettre en évidence une structure dans les réponses des agriculteurs, en résumant l'information par 9 variables synthétiques (relatives au cadre de vie, à la perception du métier, aux problèmes de l'environnement, etc.). Dans un second temps, cette analyse est complétée par une typologie en 7 groupes (intéressés par le changement, attentifs à la protection de l'environnement, adeptes de la déprise agricole, etc.). Cette démarche nous permet d'appréhender la perception de l'environnement par cette catégorie socioprofessionnelle et ainsi de mesurer l'importance accordée aux problématiques environnementales parmi les autres préoccupations actuelles.

Références

- [1] Candau, J., Deuffic, P., Ginelli, L., Lewis, N., Lyser, S., 2005, La prise en compte de l'environnement par les agriculteurs. Résultats d'enquête. CNASEA, 83 p.
- [2] Chavent M., Kuentz V., Lique B., Saracco J., 2011, Clustering of variables via the PCAMIX method, International Classification Conference, St Andrews, Ecosse.
- [3] De Soete G., Carroll J.D., 1994, K-means clustering in a low-dimensional Euclidean Space, in *New Approaches in Classification and Data Analysis*, Diday E., et al. (Eds.).
- [4] Vichi M., Kiers H.A.L., 2001, Factorial k-means analysis for two-way data, *Computational Statistics and Data Analysis*, vol 37, n° 1, p 49-64.